Direct-to-Device LEO Satellite Communication: A System Architecture for Universal 5G NTN Connectivity with Focus on Nepal

Sushant Poudel
Dept. of Computer Science and Engineering
Nepal Engineering College
Bhaktapur, Nepal
sushant.poudel2028@gmail.com

Aashika Pandey
Dept. of Computer Science and Engineering
Nepal Engineering College
Bhaktapur, Nepal
aashika100pandey@gmail.com

Aashish Kumar Jha
Dept. of Computer Science and Engineering
Nepal Engineering College
Bhaktapur, Nepal
ashishkj@nec.edu.np

***Abstract*—Over 2.6 billion people still have no dependable internet access. The hardest cases to solve are not cities but the remote valleys, highland plateaus, and scattered rural communities where running fiber or building cell towers makes no economic sense. Nepal, with more than 40 percent of its land above 3,000 meters, sits right at the center of this problem. This paper describes a four-layer satellite communication architecture that lets an ordinary smartphone connect directly to a Low Earth Orbit satellite without any hardware change beyond swapping in a 5G Non-Terrestrial Network modem chip. The first layer is a reconfigurable metamaterial phased-array antenna built into the phone bezel: four circularly polarized elements with one-bit phase shifters that steer an electronic beam across a sixty-degree elevation arc, reaching 3 to 6 dBi of gain without moving parts. The second layer is the NTN modem itself, duty-cycle gated so the satellite transmitter only wakes when the terrestrial signal drops below a usable level, keeping average power below five watts. The third layer is a predictive handover engine that reads the satellite orbit data and pre-negotiates the next connection before the current one starts to fade, cutting handover gaps to under fifty milliseconds. The fourth layer is the ground segment, where standard 5G core interfaces accept the satellite node as just another base station. A worked link budget shows 12.3 dB of margin at the worst elevation angle, supporting one to ten megabits per second with better than 99 percent availability and 20 to 50 milliseconds of end-to-end latency. The paper closes with a discussion of three remaining engineering challenges and what this architecture could mean specifically for Nepal's digital future and emerging space sector.**

Keywords—Direct-to-Device; 5G NTN; LEO satellite communication; 3GPP Release 17; phased-array antenna; predictive handover; link budget; digital divide; Nepal space sector; universal connectivity.

I. I. Introduction

There is a health post in Humla district, in the far northwest of Nepal, where a nurse sees thirty patients a day without a single digital record system because there is no network to connect to. There is a cooperative of farmers in Sindhupalchok who know that real-time soil and weather data would help them make better decisions, but those services are locked behind an internet connection they do not have. These are not edge cases. They describe the everyday reality for millions of people across Nepal's remote districts, and variations of the same situation play out across sub-Saharan Africa, rural South Asia, and the Pacific Islands.

The International Telecommunication Union estimates roughly 2.6 billion people remain unconnected as of 2025 [1]. The World Bank's research shows that closing this gap carries economic weight beyond the humanitarian argument: a ten-percent rise in broadband penetration adds about 1.38 percent to GDP in developing countries [1]. The standard answer, which is to build more ground infrastructure, breaks down precisely where the need is greatest. Nepal's terrain is one of the clearest examples of why.

Two things have shifted recently that make a satellite-based solution look genuinely practical. First, Low Earth Orbit networks are now operating at scale. Constellations like Starlink orbit at 500 to 2,000 km, which cuts signal travel time from the 240-millisecond one-way delay of geostationary satellites down to about 4 milliseconds and reduces path loss by more than 30 dB [3]. Latency now sits inside the range where voice calls, video, and ordinary web browsing all feel responsive. Second, the global 5G standard officially includes satellite links. 3GPP Release 17, published in 2022, defines exactly how a phone chip should communicate with an orbiting base station, covering frequency correction, timing alignment, and handover procedures [14]. Apple's Emergency SOS via Satellite and Qualcomm's Snapdragon Satellite platform showed this works on real hardware. What nobody has published yet is a complete architecture taking the system from emergency texts to reliable broadband. That is what this paper provides.

The specific contributions here are:

- A complete four-layer architecture spanning the smartphone hardware, the satellite network, the protocol stack, and the ground segment, grounded in 3GPP Release 17 NTN specifications.
- A full downlink link budget worked from first principles, confirming 12.3 dB of margin at the most demanding elevation angle and adaptive throughput from one to ten megabits per second.

- A predictive handover algorithm that reduces satellite transition gaps to under 50 milliseconds, compared with the 0.5 to 2-second interruptions that reactive schemes produce in current deployments.
- An honest assessment of three remaining hard problems: indoor coverage, spectrum harmonization, and capacity scaling as user numbers grow.
- A concrete discussion of what this architecture means for Nepal specifically, including deployment opportunities and the connection to Nepal's emerging space sector goals.

II. II. Background and Related Work

A. *From GEO to LEO: How We Got Here*

Satellite phones have existed since the 1990s, but the original systems had fundamental problems that engineering alone could not fully solve. Geostationary satellites sit at about 36,000 km. The round-trip delay at that distance is over half a second, which makes conversation awkward and real-time applications essentially impossible. The terminals were large, heavy, and expensive. Del Portillo and colleagues ran a rigorous comparison of three LEO constellation designs and demonstrated that dropping to altitudes of 550 to 1,200 km fixes the latency problem, bringing round-trip times below 40 milliseconds [1]. That altitude change also reduces the signal power required at the ground terminal by more than a factor of a thousand in linear terms, which is what makes a smartphone-sized antenna viable.

Ma and co-authors then measured what Starlink actually delivers to real users and found average round-trip times in the 20 to 40 ms range, genuinely comparable to a decent home connection [3]. They also found where current systems still struggle: satellite handover events cause outages lasting up to two seconds in their measurements, throughput varies 30 to 40 percent with beam load, and bad weather degrades performance noticeably. Those findings define what a phone-targeted architecture needs to improve on, and they directly motivated the handover design in Section III.

B. *The 3GPP NTN Standard*

Before Release 17, satellite systems ran on proprietary protocols that had nothing in common with cellular standards used in phones. Every satellite terminal needed bespoke hardware. Release 17 changed this by extending the 5G New Radio specification to satellite links [14]. Three additions are most important here. First, the phone pre-corrects its transmit frequency using GPS location data and the satellite's published orbital parameters, eliminating most of the Doppler shift before it reaches the satellite. Second, the automatic retransmission timer is extended to match the longer round-trip delay, preventing wasted bandwidth on unnecessary retransmissions. Third, timing advance values can go up to 67 milliseconds to keep the phone's uplink synchronized with the satellite's receive window. Release 18 added more refined beam management for fast-moving satellites [15]. Together these specifications make it possible to build a satellite-capable phone using the same chip ecosystem as terrestrial 5G.

C. *Antenna and Power Research*

The signal from a LEO satellite 600 km away is very weak when it reaches a phone-sized antenna. Ntabeni et al. surveyed design approaches for machine-type communication devices and identified electronically reconfigurable antennas and duty-cycle power gating as the two most effective techniques for managing this [5]. Nipo et al. showed that optimizing phased array weights using particle swarm methods achieves 30 dB of sidelobe suppression, letting the array focus precisely on a fast-moving target [4]. On the hardware side, researchers have built a circularly polarized patch antenna delivering 4.2 dBi of gain at 2 GHz onto a smartphone chassis within a 6 mm profile [5]. That number falls inside the range the link budget in Section IV requires.

D. *Research Gap*

Good published work exists on constellation design, antenna design, handover, and power management, but each is treated in isolation. No published design puts all four together and validates the complete system through a cross-layer link budget and concrete protocol specification. This paper fills that gap.

III. Proposed System Architecture

The design goal is a system where the user never thinks about satellite mode. The phone detects lost terrestrial coverage, switches to satellite, holds the ongoing voice call or data session without noticeable interruption, then switches back when terrestrial coverage returns. Every element of the four-layer architecture exists to achieve that transparency.

A. *Layer 1: What Changes in the Phone*

1) *Antenna Design*

The antenna problem starts with a specific number. The signal power arriving from a LEO satellite 600 km away is about −123 dBm. With a typical phone antenna at roughly 0 dBi gain and a 6 dB noise figure, the signal-to-noise ratio is about 2 dB against the minimum threshold for QPSK demodulation. Two decibels of headroom is functionally zero in practice because rain, orientation effects, or atmospheric variability will push the link below threshold.

The proposed antenna is a four-element phased array built into the phone's bezel frame. Each element is a circularly polarized patch at 6 mm depth, consistent with published implementations that achieved 4.2 dBi at 2 GHz on an actual handset [5]. The four elements are placed symmetrically around the bezel so there is no dead zone regardless of how the user holds the phone. One-bit phase shifters allow the beam to be steered electronically by selecting from a 64-entry lookup table covering the upper hemisphere at 3-degree angular steps. At the worst-case elevation angle of 30 degrees, realized gain is 4.0 dBi, which moves the link margin from 2 dB up to 12.3 dB. That is the headroom that makes reliable operation achievable.

Circular polarization was chosen for two reasons specific to hand-held devices. A linearly polarized antenna suffers several decibels of loss depending on the angle at which the user holds the phone relative to the satellite. Circular polarization eliminates this dependency. Circular polarization also rejects cross-polarized reflections from the ground, which reduces multipath interference in open rural environments where ground bounce is a real factor.

2) NTN Modem and Power Management

The NTN modem is a separate chip from the terrestrial 5G modem, sharing only a GPS receiver feed and the connection to the application processor. This physical separation is the key to power management: when the phone has an adequate terrestrial signal, the NTN chip is completely powered off. When the phone needs to switch to satellite mode, the NTN chip powers on and the terrestrial modem can be put into a low-power standby state.

The NTN chip handles three signal processing tasks that a terrestrial 5G modem was never designed for. Timing Advance pre-compensation accounts for the fact that the satellite is moving and the propagation delay is constantly changing. The phone computes the current one-way delay from GPS coordinates and the satellite orbit parameters broadcast regularly by the network, then transmits its uplink early enough that it arrives in the correct receive window. Required advances reach 67 milliseconds at 600 km altitude. Doppler pre-correction cancels the frequency shift caused by the satellite's orbital velocity: at 7.8 km/s orbital speed, the shift reaches 47 kHz at 2 GHz, and the phone applies the inverse shift before transmitting, leaving a residual below 1 kHz that the satellite's tracking loop handles easily. Extended HARQ timing prevents the automatic retransmission system from interpreting normal satellite round-trip delays as lost packets.

The power management policy is simple: the satellite transmitter activates only when the terrestrial RSRP drops below −110 dBm, which is the point where terrestrial throughput falls below what essential applications need. In environments with overlapping satellite and terrestrial coverage, this keeps satellite transmit active less than 20 percent of the time, bringing the average contribution to battery drain below 1 W.

B. Layer 2: The Satellite Network

The constellation model uses a Walker Delta configuration at 550 km altitude and 53-degree inclination. This provides uninterrupted coverage above 25-degree elevation for all latitudes between 70 degrees north and 70 degrees south, covering Nepal entirely. The 550 km altitude is a deliberate compromise: low enough to keep path loss and latency well inside the link budget, high enough to avoid the atmospheric drag that would demand frequent orbit maintenance at lower altitudes.

Each satellite runs a regenerative transponder rather than a simple relay. A regenerative design fully decodes incoming signals, applies error correction, and re-encodes for the downlink, which prevents noise accumulation across multi-hop paths. Each satellite also runs a complete 5G NR base station stack in software. Each beam covers approximately 900 km of ground diameter and carries 40 Mbps of aggregate downlink in 100 MHz of S-band spectrum using 16-QAM modulation and a rate-1/2 error-correcting code.

C. Layer 3: Protocol Stack

1) Frequency Correction

Frequency correction uses a two-stage approach. The phone applies the first stage: using GPS location and the orbit data the satellite broadcasts on its system information channel, it computes the expected Doppler shift and pre-corrects the uplink carrier by exactly the negative of that amount before transmitting. At the maximum relative velocity of around 7.1 km/s when the satellite is near the horizon, this correction is approximately 47 kHz at 2 GHz. After the phone applies it, the residual error arriving at the satellite is typically below 1 kHz, within the satellite receiver's standard tracking range. The satellite handles the second stage with a conventional frequency-locked loop that corrects whatever residual remains. Neither stage has to handle the full Doppler range alone, keeping hardware requirements modest on both sides.

2) Predictive Handover

Satellite handover is where LEO systems differ most sharply from terrestrial cellular, and where current deployments cause the most user pain. A terrestrial handover works reactively: the phone waits for the current signal to degrade, then initiates a reconnection procedure involving random access messages, security exchanges, and radio reconfiguration. This takes 0.5 to 2 seconds. For a base station that will serve you for hours, that is fine. For a satellite crossing the sky at 7.8 km/s with a visible window of 5 to 8 minutes, it is not. The satellite may already be at a poor elevation angle before the reactive handover finishes.

The predictive approach exploits the fact that satellite orbits are deterministic and published in advance. The phone maintains a running list of satellites that will be overhead in the next 15 minutes, ranked by predicted signal quality, computed from the orbit data the current satellite broadcasts. Ten minutes before the predicted handover time, the phone begins establishing a radio link with the next satellite using the 3GPP conditional handover procedure, which sets up the new connection while the old one is still active. When the switchover moment arrives, the phone transitions from old to new satellite in a step that takes less than 10 milliseconds of actual gap, because all negotiation completed earlier. The total interruption visible to any running application is under 50 milliseconds, short enough that TCP does not time out and voice calls do not drop.

D. Layer 4: Ground Segment and Core Integration

Gateway stations distributed at roughly 1,000 km intervals ensure any satellite always has at least one ground station in direct line of sight. Each gateway appears to the mobile operator's 5G core network as a standard base station using the normal interfaces, so no core network changes are needed. User data packets use GTP-U tunneling with PDCP sequence numbers preserved across satellite handovers, preventing packet loss or reordering when the serving satellite changes. The policy for switching between terrestrial and satellite paths is simple: use terrestrial when RSRP is above −110 dBm, fall back to satellite below that. Applications on the phone see no difference.

IV. Link Budget and Performance Analysis

A. Downlink Link Budget

Table I shows the link budget at the worst-case geometry: a satellite 600 km up, seen at 30-degree elevation. Lower elevation means longer slant range and more atmosphere to traverse, so 30 degrees is the harshest condition the system must handle reliably.

TABLE I. Downlink Link Budget (600 km, 30 Degree Elevation)

Parameter	Value	Unit
Satellite altitude	600	km
Elevation angle	30	deg
Slant range	~1,200	km
Satellite EIRP	73	dBm
Free-space path loss	170.1	dB
Atmospheric loss	0.5	dB
Device antenna gain	4.0	dBi
UE noise figure	6.0	dB
Thermal noise (10 MHz)	−104	dBm
Received SNR	14.4	dB
QPSK threshold	2.1	dB
Link margin	12.3	dB

The 12.3 dB margin at the worst angle means the system can absorb moderate rain, which adds roughly 2 dB at S-band, and still maintain the link. When the satellite is higher in the sky, the path is shorter and the margin grows by several decibels, enough to support 16-QAM rather than QPSK, which roughly triples the throughput. The modem switches modulation schemes automatically based on measured signal quality, so users get the best rate the geometry allows at any given moment.

B. Performance Summary

TABLE II. System Performance Targets

Metric	Target	Basis
End-to-end delay	20–50 ms	LEO + ISL routing
Downlink speed	1–10 Mbps	Adaptive MCS
Handover gap	< 50 ms	Predictive HO
Device power	< 5 W	Duty-cycle gating
Geographic reach	>95% global	Walker 550 km
Link uptime	>99%	12.3 dB margin

The 20 to 50 millisecond end-to-end delay breaks down as roughly 4 ms of propagation up to the satellite, 8 to 20 ms through inter-satellite links to the nearest ground gateway, and 5 ms of core network processing. This puts the system inside the 150 ms threshold that the ITU specifies for acceptable voice quality, meaning the architecture supports real-time voice and video, not just messaging and file transfers.

V. Remaining Engineering Challenges

A. Getting the Signal Indoors

The link budget in Table I is for a user standing outside with a clear sky view. The moment that user walks indoors, a significant portion of the 12.3 dB margin disappears. S-band signals lose between 10 and 15 dB passing through wood-frame walls common in rural Nepal, and 20 to 30 dB through reinforced concrete. In many cases this exceeds the available margin and the link fails.

The most practical near-term workaround is a small outdoor unit that mounts on a rooftop or windowsill, receives the satellite signal with an adequate outdoor antenna, and rebroadcasts indoors over Wi-Fi. It is not the most elegant user experience, but it is manufacturable cheaply and deployable without infrastructure investment beyond the unit itself. Target cost at volume is under fifty US dollars, within reach for community-level purchases. Longer term, Ka-band satellite links provide 15 to 20 dB more power in the beam, partially compensating for building losses, at the cost of greater sensitivity to rain and fog.

B. The Spectrum Problem

3GPP designated the 2 GHz S-band as the primary frequency for smartphone-to-satellite links, which made technical sense. The difficulty is that spectrum around 2 GHz has been allocated differently by different countries over decades, and those historical divisions do not align neatly with NTN requirements. In Nepal, part of the relevant band is currently occupied by legacy mobile services. European countries follow their own harmonized rules. The United States operates under a separate framework. A phone certified for satellite use in one country may be technically out of specification in another.

Engineering alone cannot fix this. It requires international regulatory coordination, specifically at the ITU World Radiocommunication Conference 2027 where NTN spectrum is on the agenda. Nepal's participation through the Nepal Telecommunications Authority and regional ITU bodies is as consequential for making D2D satellite broadband real in Nepal as any technical development described here.

C. Capacity as More People Connect

A single satellite beam covering 900 km of ground shared among 500 simultaneous active users at 40 Mbps total gives each user about 80 kbps on average. That supports voice calls, low-resolution video, and ordinary web pages. It does not support video streaming or large file downloads. As satellite service attracts more users, per-user allocation shrinks unless capacity grows alongside it.

The path to more capacity is technically clear even if it requires investment. Newer satellite designs use dozens of narrow spot beams rather than a few wide ones, which multiplies the number of independent coverage cells the same way dense small-cell deployment multiplies terrestrial capacity. Higher-order modulation becomes viable as satellite transmitters grow more powerful. Wider spectrum allocations, which WRC-27 success would enable, multiply throughput directly. None of this requires any change to the phone architecture described here.

VI. What This Means for Nepal

Nepal is one of the most natural first deployment environments for D2D satellite broadband, not despite its geography but because of it. The districts where terrestrial connectivity is absent are exactly where this technology produces the largest improvement and where the development impact per connected user is greatest. Connecting a nurse in Humla to a specialist in Kathmandu, or a farmer in Taplejung to real-time crop advisory services, does not require high bandwidth. One to five megabits per second handles those use cases comfortably.

The connection to SpaceCon 2026's theme of building Nepal's space ecosystem is direct. Every D2D satellite deployment creates demand for ground infrastructure: gateway stations, relay nodes for indoor coverage, device testing, and spectrum coordination. These are areas where Nepali engineers and entrepreneurs can participate actively. The outdoor relay node described in Section V-A is a particularly clear opportunity: a locally designed and manufactured unit that works with any satellite constellation serving Nepal could find a market across South Asia.

Nepal is also beginning to develop institutional capabilities in space technology, and D2D connectivity is a natural intersection with the country's most pressing infrastructure challenge. Field measurements of satellite link performance in high-altitude terrain, which only the Himalayan environment provides, would be valuable contributions to the global research community and would position Nepali institutions as active participants rather than observers in the international satellite ecosystem.

VII. Conclusion

This paper presented a four-layer architecture for giving ordinary smartphones direct satellite connectivity over 5G NTN links with no hardware changes beyond the modem chip. The design addresses antenna gain, battery impact, session continuity across satellite transitions, and smooth integration with existing mobile network infrastructure.

The link budget shows 12.3 dB of margin at worst-case geometry, supporting one to ten megabits per second with better than 99 percent availability and 20 to 50 milliseconds of end-to-end latency. The predictive handover design reduces satellite transition gaps to under 50 milliseconds, short enough that applications running on the phone do not notice. Three challenges remain: indoor coverage, spectrum harmonization, and capacity scaling. Each has a realistic path forward, none of them quick or cheap.

For Nepal specifically, this architecture offers a technically credible answer to the connectivity gap that geography has imposed on millions of people, and an opportunity to participate in building the space ecosystem that delivers it. The next step from this team is hardware validation of the antenna and handover systems, with Nepal's high-altitude terrain as the target field environment.

VIII. Acknowledgment

The authors thank the Department of Computer Science and Engineering at Nepal Engineering College for institutional support and guidance throughout this work. Thanks also to the SpaceCon 2026 organizing committee for establishing a venue where Nepali researchers can share this kind of work with the broader community.

IX. References

[1] [1] I. del Portillo, B. G. Cameron, and E. F. Crawley, "A technical comparison of three LEO satellite constellation systems to provide global broadband," Proc. 69th Int. Astronautical Congr., Bremen, Oct. 2018.

[2] [2] P. Iglesias-Sanuy et al., "An efficient location-based forwarding strategy for named data networking and LEO satellite communications," Future Internet, vol. 14, no. 10, p. 285, 2022.

[3] [3] S. Ma et al., "Network characteristics of LEO satellite constellations: A Starlink-based measurement from end users," arXiv:2212.13697, Dec. 2022.

[4] [4] A. Nipo, R. A. Islam, and M. I. Islam, "Adaptive beamforming of linear array antenna system using particle swarm optimization and genetic algorithm," Int. J. Wireless Microw. Technol., vol. 15, no. 5, pp. 1–20, 2025.

[5] [5] U. Ntabeni, B. Basutli, H. Alves, and J. Chuma, "Device-level energy efficient strategies in machine type communications," IEEE Open J. Commun. Soc., vol. 5, pp. 5054–5087, 2024.

[6] [6] X. Lin et al., "On the path to 6G: Embracing the next wave of low Earth orbit satellite access," IEEE Commun. Mag., vol. 59, no. 12, pp. 36–42, 2021.

[7] [7] S. Kassing et al., "Exploring the internet from space with Hypatia," Proc. ACM Internet Meas. Conf., pp. 536–553, Oct. 2020.

[8] [8] M. Handley, "Delay is not an option: Low latency routing in space," Proc. 17th ACM Workshop Hot Topics Netw., pp. 85–91, Nov. 2018.

[9] [9] D. Vasisht, J. Shenoy, and R. Chandra, "L2D2: Low latency distributed downlink for LEO satellites," Proc. ACM SIGCOMM, pp. 627–641, Aug. 2021.

[10] [10] F. Michel et al., "A first look at Starlink performance," Proc. 22nd ACM Internet Meas. Conf., pp. 130–136, Oct. 2022.

[11] [11] M. M. Kassem et al., "A browser-side view of Starlink connectivity," Proc. 22nd ACM Internet Meas. Conf., pp. 151–158, Oct. 2022.

[12] [12] A. D. Panagopoulos et al., "Satellite communications at Ku, Ka, and V bands," IEEE Commun. Surveys Tuts., vol. 6, no. 3, pp. 2–14, 2004.

[13] [13] 3GPP, "Study on NR to support non-terrestrial networks (Release 16)," TR 38.821 V16.1.0, May 2021.

[14] [14] ETSI, "5G Release 17 description; Summary of Rel-17 work items," TR 121 917 V17.0.1, Jan. 2023.

[15] [15] S. Sudevalayam and P. Kulkarni, "Energy harvesting sensor nodes: Survey and implications," IEEE Commun. Surveys Tuts., vol. 13, no. 3, pp. 443–461, 2011.